

Robust Network Reconstruction in Polynomial Time

David Hayden, Ye Yuan and Jorge Gonçalves

Abstract—This paper presents an efficient algorithm for robust network reconstruction of Linear Time-Invariant (LTI) systems in the presence of noise, estimation errors and unmodelled nonlinearities. The method here builds on previous work [1] on robust reconstruction to provide a practical implementation with polynomial computational complexity. Following the same experimental protocol, the algorithm obtains a set of structurally-related candidate solutions spanning every level of sparsity. We prove the existence of a magnitude bound on the noise, which if satisfied, guarantees that one of these structures is the correct solution. A problem-specific model-selection procedure then selects a single solution from this set and provides a measure of confidence in that solution. Extensive simulations quantify the expected performance for different levels of noise and show that significantly more noise can be tolerated in comparison to the original method.

I. INTRODUCTION

An active topic of dynamical systems research and an ubiquitous problem is that of *network reconstruction*, or inferring information about the structural and dynamical properties of a networked system [1]–[7]. This problem is motivated by a diverse range of fields where an unknown system can be described by interconnecting subsystems acting between accessible system states. By perturbing the system and measuring these states, we seek to understand the nature and dynamics of their interactions. In general, this is an underdetermined problem and is further complicated by the fact that not all states may be available for measurement and that there may also be additional, unknown states. Here we focus on noisy LTI systems and seek to obtain reliable structural and dynamic information involving the measured states, whilst leaving the hidden states unrestricted.

One prominent example from Systems Biology is the identification of gene regulatory networks, which describe the interactions between genes via biochemical mechanisms. Despite the stochastic and nonlinear nature of biological systems, a linear-systems description provides an appealing and tractable option. Measurements of the expression levels of individual genes are routinely available (for example from microarray experiments) and a number of these are taken as system states. Additional ‘hidden’ states are required to accurately model the effect of protein and other metabolic interactions, any process with higher than first-order dynamics and any important genes that are not selected as system states. In such applications, the assumption of full-state measurement may well lead to incorrect conclusions.

Nevertheless, there are a number of published methods to obtain a representative linear model from full state measure-

ments, for example [2] using Linear Matrix Inequalities, [3] using a 1-norm residual and [4] using compressive sensing. There are many other approaches to network reconstruction, for example Bayesian [5], information theoretic [6] and Boolean [7], in all of which the solution is biased, most commonly towards sparsity, to compensate for the lack of information. The vice of all these methods is that if the true solution is not sparse, a sparse solution will be obtained anyway.

Dynamical structure functions were introduced in [8] as a means of representing the structure and dynamics of an LTI system at a resolution consistent with the number of measured states. Exactly how much additional information is required to reconstruct a network from the system transfer matrix could then be quantified. In particular, without extra information, it is possible to obtain a solution that justifies any prior assumption. Applications of dynamical structure functions include multi-agent systems [9], [10] in addition to network reconstruction [11], where necessary and sufficient conditions were given for exact reconstruction from the transfer matrix.

In [1], this approach was made robust to uncertainty in the transfer matrix by obtaining dynamical structure functions that are closest, in some sense, to the data. Since the dynamical structure function is no longer unique, it is also necessary to estimate Boolean network structure, that is, an unweighted directed graph of causal connections. The approach taken was to calculate the optimum dynamical structure function for every possible Boolean network for a given number of measured states, p , then use a model selection technique to select the best estimate. This approach suffers chiefly from a high computational complexity of $O(2^p)$, which limits its use to relatively small networks.

Here we present an extension of [1] that does not require all of the Boolean network structures to be considered. In fact, we obtain a set of $p^2 - p + 1$ candidate structures by judiciously removing links from the fully-connected structure, with complexity in the number of structures that must be considered of $O(p^3)$. This set contains one structure for every level of sparsity and hence there is no prior bias towards sparser solutions.

Section II defines dynamical structure functions and states necessary previous results. The main result of a polynomial-time reconstruction algorithm is then presented in Section III. Section IV introduces a new model selection procedure. Section V then compares the performance of two variants of the method introduced here with that of [1] in extensive random simulations. Conclusions and an outline of future work are given in Section VI.

II. DYNAMICAL STRUCTURE

We consider Linear Time-Invariant (LTI) systems of the following form:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= [I \ 0]x \end{aligned} \quad (1)$$

where $x \in \mathbb{R}^n$ is the full state vector, $u \in \mathbb{R}^m$ is the vector of inputs, $y \in \mathbb{R}^p$ is the vector of measured states, with $0 < p < n$, and I is the $p \times p$ identity matrix. That is, we assume that some of the states are directly measured and some are not. It is also possible to consider a more general form of C matrix, as in [12]. By eliminating the hidden states, the dynamical structure function representation can be derived (see [11]) as:

$$Y = QY + PU \quad (2)$$

where Y and U are the Laplace transforms of y and u and Q and P are strictly proper transfer matrices. The hollow matrix Q is the *Internal Structure* and dictates direct causal relationships between measured states, which may occur via hidden states. The matrix P is the *Control Structure* and similarly defines the relationships between inputs and measured states that are direct in the sense that they do not occur via any other measured states. The dynamical structure function is defined as (Q, P) and the Boolean dynamical structure is defined as the Boolean matrices $\mathcal{B}(Q)$, $\mathcal{B}(P)$ that have the same zero elements as Q and P .

A. Dynamical Structure Reconstruction

The problem of network reconstruction was cast in [11] as a two-stage process, whereby the transfer matrix G is first obtained from input-output data by standard system identification techniques and the dynamical structure function is then obtained from G . Here we state the main results for the second stage of this process. The dynamical structure function for a given state-space realisation is unique, and related to the transfer matrix, G , as follows:

$$G = (I - Q)^{-1}P \quad (3)$$

Whilst every (Q, P) uniquely specifies a G , there are many such possible (Q, P) s for any given G and a dynamical structure function is said to be *consistent* with a given transfer matrix if, and only if, there exists a state-space realisation for which the dynamical structure function (at the considered resolution) satisfies (3). Exact reconstruction is therefore possible if, and only if, there is only one (Q, P) that is consistent with G , which requires some *a priori* knowledge of (Q, P) . Corollary 1 defines an experimental protocol under which this condition is met.

Corollary 1 ([11]). *If G is full rank and p inputs are applied, where there are p measured states and each input directly and uniquely affects only one measured state (such that the matrix P can be made diagonal), then (Q, P) can be recovered exactly from G .*

Essentially the zero elements of P comprise sufficient knowledge to obtain (Q, P) directly from (3). Here we assume that

the conditions of the above corollary have been met, which ensures solution uniqueness when G is known perfectly.

B. Robust Dynamical Structure Reconstruction

In practice, the transfer matrix G will be a noisy estimate of the true transfer matrix, as considered in [1]. As a result, the dynamical structure function obtained directly from G may not be the best approximation of that of the true system. The noise is expressed as a perturbation on the true transfer matrix G_t , for example as feedback uncertainty: $G_t = (I + \Lambda)^{-1}G$ for some transfer matrix Λ . Since G will typically admit a fully-connected dynamical structure function, a strategy to estimate the Boolean dynamical structure consistent with G_t is required.

Consider the i^{th} Boolean dynamical structure function and denote (Q_i, P_i) as a dynamical structure function with this Boolean structure. We can relate G to a transfer matrix G_i that is consistent with (Q_i, P_i) by $G = (I + \Delta_i)G_i$ for some transfer matrix Δ_i . In this case, (3) becomes: $(I + \Delta_i)^{-1}G = (I - Q_i)^{-1}P_i$, which can be rearranged as:

$$\begin{aligned} \Delta_i &= GP_i^{-1}(I - Q_i) - I \\ &= GX_i - I \end{aligned} \quad (4)$$

where $X_i = P_i^{-1}(I - Q_i)$ and has the same non-diagonal Boolean structure as Q_i . By minimising some norm of Δ_i , with respect to X_i , we can obtain the X_i and hence the corresponding dynamical structure function that is consistent with the closest transfer matrix to G .

The approach of [1] was to minimise $\|\Delta_i\|$ for every possible Boolean Q_i (of which there are 2^{p^2-p} since Q has $p^2 - p$ degrees of freedom), then use Akaike's Information Criterion (AIC) [13] to select a solution by penalising the number of nonzero elements in Q . Specifically, let \mathcal{X}_i be the set of all X that satisfy the constraints of the i^{th} Boolean X , and minimise the Frobenius norm over $s = j\omega$ of Δ_i as follows:

$$\delta_i^2 = \inf_{X \in \mathcal{X}_i} \|GX - I\|_F^2 \quad (5)$$

to obtain a measure of the smallest distance δ_i from G to G_i . This choice of norm allows the problem to be cast as a least squares optimisation, and we denote this method M_0 .

Algorithm M_0

```

for  $i = 1 \rightarrow 2^{p^2-p}$  do
   $\delta_i^2 = \inf_{X \in \mathcal{X}_i} \|GX - I\|_F^2$ 
   $X_i = \arg \inf_{X \in \mathcal{X}_i} \|GX - I\|_F^2$ 
end for
Apply AIC to the set  $\{X_i\}$ .

```

Hence every Boolean structure $\mathcal{B}(X_i)$ can be associated with a distance measure δ_i from (5) and a dynamic structure X_i , which is the corresponding minimising argument. The principal problem with this approach is that the computational complexity is dominated by the number of optimisations that must be performed, which can be reduced to $p2^p$ by performing the optimisation of the columns of X separately.

III. MAIN RESULT

A. A Polynomial Time Algorithm

Here we propose an algorithm with polynomial complexity to estimate the dynamical structure function of G_t , under the conditions of Corollary 1. First, an iterative procedure is used to obtain a set \mathbb{S} , containing Boolean internal structures, with one structure (S^j) for each level of sparsity (j links). Then a model selection procedure is applied to this reduced set to select a single solution. This method is denoted M_2 and defined as follows:

Algorithm M_2

Set S^{p^2-p} as the fully-connected structure.
for $j = p^2 - p \rightarrow 1$ **do**
 Remove one link of S^j at a time to obtain a set of j structures with $j - 1$ links and calculate δ^{j-1} for each of these structures.
 Set S^{j-1} as the minimum- δ^{j-1} structure.
end for
 Set S^0 as the decoupled structure.
 Apply a model selection procedure to the set $\mathbb{S} = \{S^j\}$.

Fig. 1 illustrates this procedure for a $p = 3$ example. The number of structures that must be considered is exactly $\frac{1}{2}(p^4 - 2p^3 + 2p^2 - p) + 1$. By optimising the columns of X separately, the overall number of optimisations that must be performed is of the order $O(p^3)$, and since the complexity of the optimisations is also polynomial, the overall computational complexity is polynomial.

The reasoning behind this approach is that the fully-connected structure is composed of all the links belonging to the true structure plus extra links that afford it a smaller δ by better modelling the noise. It is intuitive that if the level of noise is not too high, removing false links should have a smaller effect on δ than removing true links, in which case all the false links would be removed first and the true structure would be encountered. In fact, we will show that if the noise is sufficiently small (in some norm) then this is always the case.

B. An Intermediate Algorithm

We note that an intermediate algorithm can be defined, with the same two-stage approach of M_2 but making use of all 2^{p^2-p} Boolean structures and hence with exponential complexity. Rather than obtain the set \mathbb{S} iteratively, this method obtains a set $\mathbb{T} = \{T^j\}$, where each T^j is the minimum- δ Boolean structure over all Boolean structures with j links.

Note that the set \mathbb{T} contains the minimum-AIC structures for each level of sparsity, and hence Method M_0 is only able to select a solution from this set. If the true Boolean structure is not in the set \mathbb{T} , then neither M_0 nor M_1 can obtain the correct structure, from which we can define:

Definition 1 (M_1 Solvability). *A given reconstruction problem is solvable by Method M_1 (and M_0) if the true Boolean internal structure is in the set \mathbb{T} .*

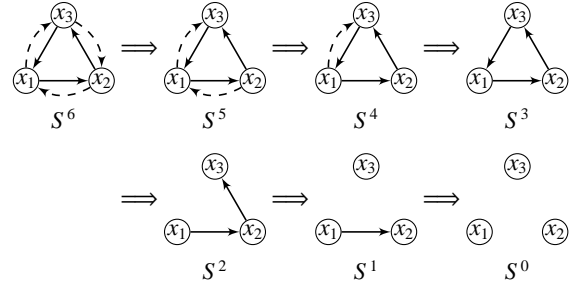


Fig. 1: An example set \mathbb{S} for Method M_2 for a network with three measured states. Solid arrows denote true links and dashed arrows denote false links. Starting with the fully-connected structure (S^6), one network is obtained for each level of sparsity by successively removing the link which results in the smallest increase in δ . A model selection technique is then used to select a single solution from this set. In this example, the true network is S^3 .

Algorithm M_1

for $j = 0 \rightarrow p^2 - p$ **do**
 Compute δ^j for all Boolean internal structures with j links.
 Set T^j as the minimum- δ^j structure.
end for
 Apply a model selection procedure to the set $\mathbb{T} = \{T^j\}$.

Therefore, if a problem is M_1 solvable, Method M_1 will always obtain a set \mathbb{T} that contains the true Boolean structure. This step has separated the uncertainty inherent in the problem due to noise from that due to the model selection process. If the model selection stage is not correct but the problem is M_1 solvable, then we have a relatively small set of candidate structures, one of which is the true structure.

In summary, M_0 considers all 2^{p^2-p} Boolean structures and selects a single solution; M_1 considers all 2^{p^2-p} Boolean structures, selects the best-fitting structure for each number of links to form a subset of $p^2 - p + 1$ structures and then selects a single solution from this subset; M_2 iteratively finds a set of $p^2 - p + 1$ Boolean structures and selects a single solution from this set, without having to consider all possible structures. We will next consider a sufficient condition under which a problem can be solved by M_2 , and it will be seen that this condition is also sufficient for the problem to be M_1 solvable.

C. Solvability Conditions for M_2

We define the solvability of M_2 as follows:

Definition 2 (M_2 Solvability). *A given reconstruction problem is solvable by Method M_2 if the true Boolean internal structure is in the set \mathbb{S} .*

First it is noted that the false links of the fully-connected structure can be removed in any order to obtain the true structure, so the path to the true structure is not unique. A set of *allowable* Boolean structures can be defined as all those that contain at least all of the true links. All other structures with no fewer links than the true structure are *non-*

allowable, as to encounter any one of these will mean that the true structure will not be reached. Fig. 2 shows δ values plotted against number of links for an example allowable set plus valid ‘paths’ which may be taken between structures by removing one link. The following Lemma provides a sufficient condition for M_2 solvability:

Lemma 1. *A given reconstruction problem is solvable by Method M_2 if all allowable structures have smaller δ than all non-allowable structures with the same number of links.*

Proof. For every allowable structure that is not the true structure, another allowable structure can always be obtained by removing one link, so a path always exists to another allowable structure. If the condition of the Lemma holds, at every stage in the M_2 algorithm an allowable structure will be selected in preference to a non-allowable structure until the true structure is reached. ■

This is equivalent to all non-allowable structures being within the open shaded region in Fig. 2. Note that if the condition of Lemma 1 holds, this also implies that the problem is M_1 solvable. We now derive, for each Boolean structure, an expression for the deviation of the δ value in the case of noise from its nominal (no noise) value. The level of noise such that the condition of Lemma 1 is met can then be characterised.

From (5), the value of δ_i for the i^{th} Boolean structure is given by:

$$\begin{aligned} \delta_i^2 &= \inf_{X \in \mathcal{X}_i} \|GX - I\|_F^2 \\ &= \inf_{Y \in \mathcal{Y}_i} \|AY - b\|_2^2 \end{aligned} \quad (6)$$

where $A = I \otimes G$, $Y = \text{vec}(X)$ and $b = \text{vec}(I)$ where I is the $p \times p$ identity matrix, \otimes is the Kronecker product and $\text{vec}(\cdot)$ is the vectorization operator. The set \mathcal{Y}_i contains a vector $Y = \text{vec}(X)$ for each element $X \in \mathcal{X}_i$. Consider a subset of Boolean internal structures, which can be obtained from the fully-connected structure by constraining at most one element in each column of X to be zero. The following Lemma relates the δ values of these structures to that of the fully-connected structure.

Lemma 2. *For any Boolean structure $\mathcal{B}(X_i)$ with no more than one zero element in each column, the value of δ_i is related to that of the fully-connected structure as follows:*

$$\delta_i^2 = \delta_{\text{full}}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_j \left(\frac{|Y_{\text{full}}(j)|^2}{A^+(j, j)} \right) d\omega \quad (7)$$

for all elements j of Y that are constrained in the i^{th} structure. Here $Y_{\text{full}} = \text{vec}(X_{\text{full}})$ for the fully-connected structure X_{full} , which has δ value δ_{full} , $A^+ := (A'A)^{-1}$ and $'$ denotes complex conjugate transpose.

The proof is given in Appendix A and follows from writing (6) as a constrained optimisation problem. In particular, this special case applies to all structures with only one element in total constrained to be zero. We now consider directly how a perturbation on the true transfer matrix G_t affects δ_i and apply a recursive argument to make use of this special case.

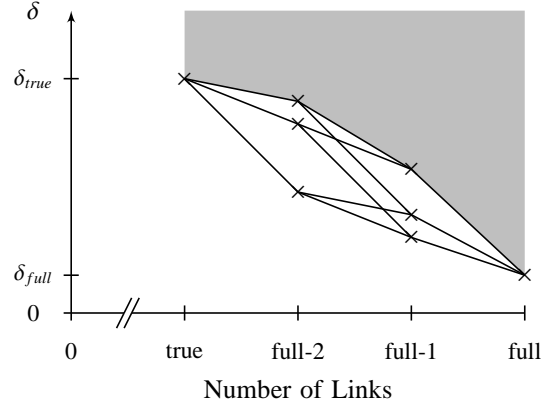


Fig. 2: An example problem showing δ values plotted against number of links for the allowable set of Method M_2 . Allowable Boolean structures (those which contain at least all of the true links) are marked by an \times . Allowable paths (those that may be taken between allowable structures by removing a single link) are shown as solid lines. The problem is M_2 solvable, by Lemma 1, if all non-allowable structures are within the open shaded region.

Write the feedback uncertainty in G_t as $\Lambda = \epsilon \Lambda_0$, where $\epsilon = \max_{i,j} \{\|\Lambda(i, j)\|_{\infty}\}$, such that $G = (I + \epsilon \Lambda_0)G_t$. Now given a perturbation $\epsilon \Lambda_0$, parameterise the solution for the δ value of the i^{th} structure by ϵ as $\delta_i(\epsilon)$. We first consider the structures obtained by constraining one link of Y_{full} .

Lemma 3. *Given an estimate of the true transfer matrix G_t as $G = (I + \epsilon \Lambda_0)G_t$, the smallest distance δ_i from G_i to G for every Boolean structure $\mathcal{B}(X_i)$ with only one zero element is given by:*

$$\delta_i^2(\epsilon) = \delta_i^2(0) + \delta_{\text{full}}^2(\epsilon) + f_i(\epsilon) \quad (8)$$

where $f_i(\epsilon)$ is continuous and satisfies $f_i(0) = 0$.

The proof is given in Appendix B and essentially consists of expressing the uncertainty in $\delta_i^2(\epsilon)$ directly using Lemma 2. For any perturbation $\epsilon \Lambda_0$ on G_t , the corresponding deviation of $\delta_i^2(\epsilon)$ from its nominal ($\epsilon = 0$) value is therefore given by $\delta_{\text{full}}^2(\epsilon) + f_i(\epsilon)$. The following Lemma extends this to describe the deviation of $\delta_i^2(\epsilon)$ for a general X_i .

Lemma 4. *Given $G = (I + \epsilon \Lambda_0)G_t$, the smallest distance δ_i from G_i to G for every Boolean structure $\mathcal{B}(X_i)$ is given by:*

$$\delta_i^2(\epsilon) = \delta_i^2(0) + \delta_{\text{full}}^2(\epsilon) + f_{ij}(\epsilon) \quad (9)$$

for any Boolean structure $\mathcal{B}(X_j)$ from which $\mathcal{B}(X_i)$ can be obtained by constraining one link. The function $f_{ij}(\epsilon)$ is continuous and satisfies $f_{ij}(0) = 0$.

The proof is given in Appendix C, where X_j is treated as the solution to a new, unconstrained problem and hence Lemma 3 can be applied. Using the results of Lemmas 3 and 4 it is possible to compute the deviation of every $\delta_i(\epsilon)$ from its nominal $\delta_i(0)$ value. By considering the difference between the deviations of any pair of allowable and non-allowable structures as a function of ϵ , it can be seen that for sufficiently small ϵ , the condition of Lemma 1 is always met and the problem is therefore M_2 solvable.

Theorem 1. Given $G = (I + \epsilon\Lambda_0)G_t$, where G_t is the true transfer matrix, there exists an $r > 0$ such that for all $\epsilon < r$, the problem is always solvable by Method M_2 . Specifically, the difference between δ_n for any non-allowable structure and δ_a for any allowable structure with the same number of links is given by:

$$\delta_n^2(\epsilon) - \delta_a^2(\epsilon) = B_n + F_{na}(\epsilon) \quad (10)$$

where the constant B_n satisfies $B_n > 0$ and $|F_{na}(\epsilon)| < B_n$ for all $\epsilon < r$.

The proof is given in Appendix D and comprises a recursive application of Lemma 4 to assess the solvability condition of Lemma 1. This result proves that every problem is M_2 solvable for sufficiently small ϵ , and given G_t and Λ_0 it is possible to calculate a bound r on the size of this ϵ . Method M_2 is then guaranteed to encounter the true structure by successively removing links from the fully-connected structure.

Remark 1 (Repeated Experiments). We note that multiple estimates of G_t are readily incorporated by simply replacing G with a block vector of these estimates.

Remark 2 (Prior Information). If something is known about the true Boolean structure, this information can easily be taken into account in Algorithm M_2 by restricting the number of structures that must be considered. This will then reduce the computational complexity.

Remark 3 (Steady-State Reconstruction). As in [1], if only steady-state measurements are available, we can reconstruct the steady-state dynamical structure function from $G(0)$. Complications arise in the case of transfer functions with zero steady-state gain, but otherwise the results of Section III can be directly applied to steady-state reconstruction. In particular, the bound of Theorem 1 is tighter and significantly easier to compute.

IV. A NEW MODEL SELECTION APPROACH

Here we introduce a method of selecting a candidate Boolean structure from the set \mathbb{S} of Method M_2 , as an alternative to AIC. Our method does not directly penalise model complexity but rather seeks to locate the correct level by identifying the subsequent loss of information as the complexity is further reduced. Denote the δ value of $S^i \in \mathbb{S}$ as δ^i and the number of links of the true structure as i_t . Note that for no noise, $\delta^i = 0$ for $i_t \leq i \leq p^2 - p$ and $\delta^i > 0$ for $i < i_t$, since the solution is unique. Now define the following normalised derivative of δ :

$$d'_i = \frac{\delta^{i-1} - \delta^i}{\delta^{i-1}}, \quad i = 1, \dots, p^2 - p \quad (11)$$

where $d'_0 := 0$ and $0 \leq d'_i \leq 1$ since $0 \leq \delta^i \leq \delta^{i-1}$. For low levels of noise, δ^i is close to zero for all $i \geq i_t$ and increases significantly for $i < i_t$, and it is this increase in δ^i that we seek to detect. In fact, as the noise level approaches zero, $d'_i \rightarrow 1$ for $i = i_t$. Taking a form of second derivative,

defined as follows, was found heuristically to improve the distinction of the true structure:

$$d''_i = \max\{d'_i - d'_{i+1}, 0\}, \quad i = 0, \dots, p^2 - p \quad (12)$$

where $d''_{p^2-p+1} := 0$ and $0 \leq d''_i \leq 1$. The candidate solution is then selected as S^j where $j = \arg \max_i \{d''_i\}$ and a measure of the confidence in this selection is given by $\max_i \{d''_i\} \in [0, 1]$.

Remark 4 (A Single Solution). Whilst it is useful to obtain a single solution for comparative simulations, in practice, a more prudent approach is advised. For example, the relative merit of each of the structures in the solution sets of M_1 or M_2 could be considered.

V. SIMULATIONS

Methods M_1 and M_2 introduced here were compared in simulation with M_0 from [1] on steady-state network reconstruction of a large number of linear test networks. Networks with three measured states and up to three hidden states were considered; for each of the 64 possible Boolean network structures with three measured states, 300 random, stable linear systems were generated. For each test system and for a range of noise variance from 10^{-5} – 10, three experimental estimates of $G(0)$ were obtained as $G(0) = (I + \Lambda)G_t(0)$, where G_t is the true transfer matrix and the elements of Λ were sampled from a zero-mean normal distribution. These estimates of $G(0)$ were then used by each method to attempt to obtain the correct steady-state network structure, given no other information about the true network. In total, 963,900 network reconstruction problems were considered.

Fig. 3 shows the average number of networks that were correctly identified by each method, for each level of noise, as a percentage of the total number of reconstruction problems attempted. Some variation in performance was observed between different network structures (some were harder to identify than others) but the results of Fig. 3 are representative of the performance of each of the methods. Also shown in Fig. 3 is the M_1 solvable limit, which is the percentage of reconstruction problems that were M_1 solvable.

The most important result is that the performances of M_1 and M_2 are almost identical, validating the use of M_2 for this problem class. Only 0.37% of problems could be solved by M_1 but not by M_2 , which is certainly justified by the significant reduction in computational complexity. The level of noise required for M_2 to fail is apparently similar to that required for M_1 to fail. In addition, M_1 and M_2 (both using the model selection procedure of Section IV) consistently outperform M_0 for all noise levels. For example, for a noise variance of 10^{-3} , approximately 60% of problems could be solved by M_0 , whereas 90% could be solved by M_1 and M_2 .

Figure 4 shows the percentage of reconstruction problems for which the set of possible Boolean solutions obtained by each method contained the true Boolean structure. Again the results for methods M_1 and M_2 are almost identical. For a noise variance of 10^{-2} , only approximately 30% of problems could be solved by M_0 , whereas for M_1 and M_2 we could be 90% certain that the correct Boolean structure is in the set.

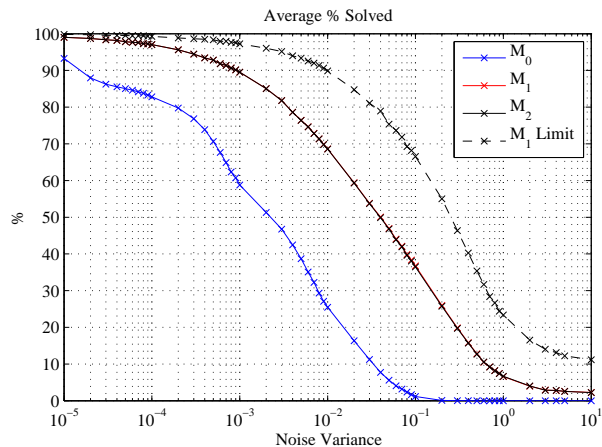


Fig. 3: Percentage of reconstruction problems successfully solved by each method. M_0 is the method from [1] with exponential complexity; M_1 is an intermediate method also with exponential complexity; M_2 is the polynomial complexity method. Also shown is the M_1 solvable limit, which is the percentage of reconstruction problems that were solvable by M_1 (and M_0). The difference between this curve and that of M_0 and M_1 is due to the incorrect solution being chosen in the model selection stage.

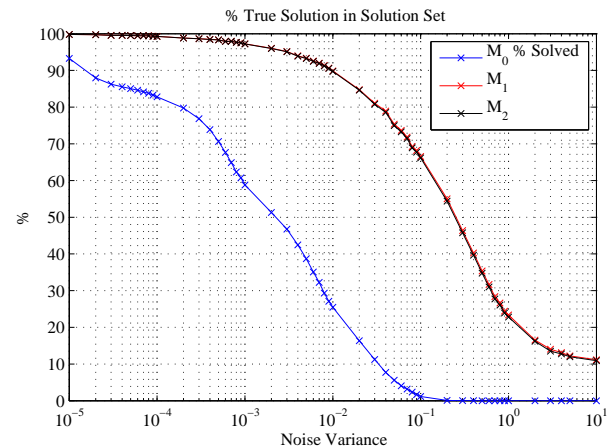


Fig. 4: Percentage of reconstruction problems where the solution set obtained by each method contained the true solution. M_0 is the method from [1] with exponential complexity; M_1 is an intermediate method also with exponential complexity; M_2 is the polynomial complexity method. Note that the solution set of method M_0 only contains a single solution.

VI. CONCLUSIONS

A. Summary

This paper introduces an algorithm with polynomial complexity that robustly reconstructs the structure and dynamics of an unknown LTI network in the presence of noise and unmodelled nonlinearities. Specifically, we estimate the dynamical structure function from a noisy estimate of the system transfer matrix. Rather than seek to obtain a sparse solution, we consider a set of solutions spanning all levels of sparsity and then select a solution from this set. Following a certain experimental protocol, we prove that every such problem is solvable by our method if the magnitude of the noise is sufficiently small, where the size of this bound depends on the properties of the system in question. The expected performance of this method is assessed in simulation, which demonstrates almost no difference in performance between the exponential and polynomial complexity versions and also shows significant improvements over the previous method.

B. Future Work

Future work will consider exactly what properties of the true system contribute to the size of the noise bound for that system. This may provide insight to develop a necessary and sufficient condition for solvability, given a certain type of noise perturbation. We also seek to relax the conditions of the experimental protocol, which is another limiting factor in the size of problem that this approach can handle.

VII. ACKNOWLEDGEMENTS

This work was supported by EPSRC grants EP/I03210X/1 and EP/P505445/1. The authors would also like to thank the reviewers for their helpful comments.

REFERENCES

- [1] Ye Yuan, Guy-Bart Stan, Sean Warnick and Jorge Gonçalves, Robust dynamical network structure reconstruction, *Automatica*, vol. 47, 2011, pp. 1230-1235.
- [2] Francesco Montefusco, Carlo Cosentino and Francesco Amato, CORE-Net: exploiting prior knowledge and preferential attachment to infer biological interaction networks, *IET Systems Biology*, vol. 4, iss. 5, 2010, pp. 296-310.
- [3] Michael M. Zavlanos, Agung A. Julius, Stephen P. Boyd and George J. Pappas, Inferring stable genetic networks from steady-state data, *Automatica*, vol. 47, 2011, pp. 1113-1122.
- [4] Borhan M. Sanandaji, Tyrone L. Vincent and Michael B. Wakin, Exact topology identification of large-scale interconnected dynamical systems from compressive observations, *Proceedings of the American Control Conference*, 2011.
- [5] Nir Friedman, Kevin Murphy and Stuart Russell, Learning the structure of dynamic probabilistic networks, *UAI'98 Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 1998, pp. 139-147.
- [6] Adam A. Margolin, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Dalla Favera and Andrea Califano ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context, *BMC Bioinformatics*, vol. 7, 2006.
- [7] Shoudan Liang, Stefanie Fuhrman and Roland Somogyi REVEAL, A general reverse engineering algorithm for inference of genetic network architectures, *Pacific Symposium on Biocomputing*, 1998.
- [8] Jorge Gonçalves, Russell Howes and Sean Warnick, Dynamical structure functions for the reverse engineering of LTI networks, *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007.
- [9] Ye Yuan, Decentralised network prediction and reconstruction algorithms, *Ph.D. Thesis, Cambridge University*, 2012.
- [10] Ye Yuan and J. Gonçalves, Minimal-time network reconstruction of DTLTI system, *Proceedings of 49th IEEE Conference on Decision and Control*, 2010.
- [11] Jorge Gonçalves and Sean Warnick, Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks, *IEEE Transactions on Automatic Control*, vol. 53, 2008, pp. 1670-1674.
- [12] Enoch Yeung, Jorge Gonçalves, Henrick Sandberg and Sean Warnick, The meaning of structure in interconnected dynamic systems, to appear in *IEEE Control Systems Magazine Special Invited Issue: Designing Controls for Modern Infrastructure Networks*, 2012.
- [13] Akaike Hirotugu, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, vol. 19, 1974, pp. 716-723.

A. Proof of Lemma 2

Proof. Equation (6) can be written as a constrained optimisation problem as follows:

$$\delta_i^2 = \inf_Y \|AY - b\|_2^2 \quad \text{subject to} \quad C_i'Y = 0 \quad (13)$$

where C_i contains the relevant columns of the $p^2 \times p^2$ identity matrix. The constraints are appended to the cost function by a vector of Lagrange multipliers μ_i :

$$\delta_i^2 = \inf_Y \frac{1}{2\pi} \int_{-\infty}^{\infty} |AY - b|^2 + \mu_i' C_i' Y d\omega \quad (14)$$

$$C_i' Y = 0$$

with $s = j\omega$ for all transfer functions. Solving for Y_i and μ_i yields:

$$Y_i = (A'A)^{-1} \left(A'b - \frac{1}{2} C_i \mu_i \right) \quad (15)$$

$$\mu_i = 2 \left(C_i' (A'A)^{-1} C_i \right)^{-1} C_i' (A'A)^{-1} A'b$$

The following abbreviations are made to simplify notation: $A^+ = (A'A)^{-1}$, $D_i = C_i' A^+ C_i$ and $E_i = C_i D_i^{-1} C_i' A^+$. It is noted that the fully-connected (unconstrained) solution is $Y_{full} = A^+ A'b$ and the corresponding δ_{full} given by:

$$\delta_{full}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} b' (I - AA^+ A') b d\omega \quad (16)$$

Using (14) and (15), for the i^{th} Boolean structure the solution is $Y_i = A^+ (I - E_i) A'b$ and the optimum δ_i is given as follows:

$$\begin{aligned} \delta_i^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} b' (I - AA^+ A' + AE_i' A^+ E_i A') b d\omega \\ &= \delta_{full}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} b' AE_i' A^+ E_i A' b d\omega \\ &= \delta_{full}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} b' AA^+ C_i D_i^{-1} C_i' A^+ A' b d\omega \\ &= \delta_{full}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_{full}' C_i D_i^{-1} C_i' Y_{full} d\omega \end{aligned} \quad (17)$$

where $C_i' Y_{full}$ is the vector of the elements of Y_{full} that are constrained in the i^{th} Boolean structure. Equation (17) expresses the optimum δ value of every Boolean structure in terms of the δ value of the fully-connected structure plus the effect of constraining some of the links to be zero.

The matrix D_i is block-diagonal and composed of elements of A^+ , which is a block-diagonal matrix given by $A^+ = I \otimes (G'G)^{-1}$, where I is the $p \times p$ identity matrix. If no more than one element is constrained in each column of X , then D_i is diagonal and composed of diagonal elements of A^+ . In this special case, (17) reduces to:

$$\delta_i^2 = \delta_{full}^2 + \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_j \left(\frac{|Y_{full}(j)|^2}{A^+(j, j)} \right) d\omega \quad (18)$$

for all elements j of Y that are constrained in the i^{th} structure. ■

B. Proof of Lemma 3

Proof. Since X_i has only one zero element, Lemma 2 results in:

$$\delta_i^2(\epsilon) - \delta_{full}^2(\epsilon) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|Y_{full}(j)|^2}{A^+(j, j)} d\omega \quad (19)$$

where j is the index of the constrained element in Y_{full} . Since $G = (I + \epsilon \Lambda_0) G_t$, A can be written as $A = (I + (I \otimes \epsilon \Lambda_0)) A_t$, where $A_t = I \otimes G_t$. Due to the equivalence of dynamic uncertainty models it is possible to parameterise the uncertainty in a number of forms, which allows us to write: $(A'A)^{-1} = (A_t' A_t)^{-1} (I + \Lambda_1(\epsilon))$ for some $\Lambda_1(\epsilon)$ which is also a function of Λ_0 and A_t . Similarly we can write: $(A'A)^{-1} A' = (A_t' A_t)^{-1} A_t' (I + \Lambda_2(\epsilon))$ for some other $\Lambda_2(\epsilon)$.

Denoting $A_t^+ = (A_t' A_t)^{-1}$, $\bar{A}_t = A_t^+ A_t'$ and $\bar{A} = A^+ A'$, the integrand of (19) is given by:

$$\begin{aligned} \frac{|Y_{full}(j)|^2}{A^+(j, j)} &= \frac{|\bar{A}(j, :) b|^2}{A^+(j, j)} \\ &= \frac{|\bar{A}_t(j, :)(I + \Lambda_1)b|^2}{A_t^+(j, :)(I + \Lambda_2(\epsilon))} \\ &= \frac{|\bar{A}_t(j, :)|^2}{A_t^+(j, j)} (1 - \alpha_j(\epsilon)) + \beta_j(\epsilon) \end{aligned} \quad (20)$$

where $(j, :)$ and $(:, j)$ index respectively the j^{th} row and column of a matrix, and the dependence of Λ_1 and Λ_2 on ϵ has been omitted. The functions $\alpha_j(\epsilon)$ and $\beta_j(\epsilon)$ are continuous functions of ϵ , satisfying $\alpha_j(0) = 0$ and $\beta_j(0) = 0$. Note that $\delta_i^2(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\bar{A}_t(j, :)|^2}{A_t^+(j, j)} d\omega$ and hence (19) can be written:

$$\delta_i^2(\epsilon) - \delta_{full}^2(\epsilon) = \delta_i^2(0) + f_i(\epsilon) \quad (21)$$

where $f_i(\epsilon) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \beta_j(\epsilon) - \frac{|\bar{A}_t(j, :)|^2}{A_t^+(j, j)} \alpha_j(\epsilon) d\omega$ is also a continuous function of ϵ and satisfies $f_i(0) = 0$. ■

C. Proof of Lemma 4

Proof. Define \tilde{Y}_j as Y with every element that is constrained to be zero in the j^{th} Boolean structure removed. Similarly \tilde{A}_j is obtained from A by removing the columns of A corresponding to the constrained elements of Y_j . Taking \tilde{Y}_j as the unconstrained minimising argument of $\delta_j^2 = \inf_{\tilde{Y}_j} \|\tilde{A}_j \tilde{Y}_j - b\|_2^2$, Lemma 3 gives:

$$\delta_i^2(\epsilon) - \delta_j^2(\epsilon) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\tilde{Y}_j(k)|^2}{\tilde{A}_j^+(k, k)} d\omega \quad (22)$$

where k is the index of the constrained element in $\tilde{Y}_j = \tilde{A}_j^+ \tilde{A}_j' b$. Recall that $A_t = I \otimes G_t$ and obtain \tilde{A}_t from A_t by removing the columns of A_t corresponding to the constrained elements of Y_j . Then we can write $\tilde{A}_j = (I + (I \otimes \epsilon \Lambda_0)) \tilde{A}_t$ and this problem is now in the same form as that of Lemma 3 and the proof follows as before. ■

D. Proof of Theorem 1

Proof. A problem is M_2 solvable from Lemma 1 if for every allowable structure Y_a and every non-allowable structure Y_n with the same number of links, $\delta_n^2(\epsilon) > \delta_a^2(\epsilon)$. From Lemma 4, for every allowable structure Y_a , $\delta_a^2(\epsilon)$ is given by:

$$\begin{aligned}\delta_a^2(\epsilon) &= \delta_{a_1}^2(\epsilon) + f_{a,a_1}(\epsilon) \\ &= \delta_{full}^2(\epsilon) + f_{a,a_1}(\epsilon) + \sum_i f_{a_i,a_{i+1}}(\epsilon) \\ &= \delta_{full}^2(\epsilon) + F_a(\epsilon)\end{aligned}\quad (23)$$

where a_i denotes an allowable structure with i links more than Y_a , from which Y_a can be obtained by removing links. Here we have used the facts that any allowable structure can be obtained by removing one link from another allowable structure and that $\delta_a(0) = 0$ for all allowable structures due to solution uniqueness from Corollary 1. From Lemma 4, the functions $f_{a_i,a_{i+1}}(\epsilon)$ are continuous and satisfy $f_{a_i,a_{i+1}}(0) = 0$, hence $F_a(\epsilon)$ (which is a sum of these functions) also satisfies these properties.

Similarly, for every non-allowable structure Y_n :

$$\begin{aligned}\delta_n^2(\epsilon) &= \delta_n^2(0) + \delta_{n_1}^2(\epsilon) + f_{n,n_1}(\epsilon) \\ &= \sum_i \delta_{n_i}^2(0) + \delta_{full}^2(\epsilon) + \sum_i f_{n_i,n_{i+1}}(\epsilon) \\ &= B_n + \delta_{full}^2(\epsilon) + F_n(\epsilon)\end{aligned}\quad (24)$$

where n_i denotes a non-allowable structure with i links more than Y_n , from which Y_n can be obtained by removing links. The constant $B_n = \sum_i \delta_{n_i}^2(0) > 0$ again due to solution uniqueness from Corollary 1 and, as in the allowable case, $F_n(\epsilon)$ is continuous and satisfies $F_n(0) = 0$. The solvability condition is then:

$$\delta_n^2(\epsilon) - \delta_a^2(\epsilon) = B_n + F_n(\epsilon) - F_a(\epsilon) > 0 \quad (25)$$

where the difference between δ_n and δ_a is characterised by a perturbation from each of their nominal ($\epsilon = 0$) values. From the properties of F_n and F_a , the function $F_{na} = F_n - F_a$ is also continuous and satisfies $F_{na}(0) = 0$ and hence there always exists some $r_{na} > 0$ such that $|F_{na}(\epsilon)| < B_n$ for all $\epsilon < r_{na}$. The solvability condition of (25) is then satisfied as follows:

$$\begin{aligned}\delta_n^2(\epsilon) - \delta_a^2(\epsilon) &= B_n + F_{na}(\epsilon) \\ &\geq B_n - |F_{na}(\epsilon)| \\ &> 0\end{aligned}\quad (26)$$

for all $\epsilon < r_{na}$. Taking r to be the minimum r_{na} over all pairs of allowable and non-allowable structures will therefore ensure the problem is M_2 solvable for all $\epsilon < r$. ■